(2)

A CONNECTIONIST/CONTROL ARCHITECTURE FOR WORKING MEMORY AND WORKLOAD: *WHY WORKING MEMORY IS NOT 7+ /-2*

Technical Report AIP - 26

Mark Detweiler & Walter Schneider

Learning Research and Development Center
Univerity of Pittsburgh
Pittsburgh, PA 15260

The Artificial Intelligence and Psychology Project

Departments of
Computer Science and Psychology
Carnegie Mellon University

Learning Research and Development Center
University of Pittsburgh

A CONNECTIONIST/CONTROL ARCHITECTURE FOR WORKING MEMORY AND WORKLOAD: WHY WORKING MEMORY IS NOT 7+ /-2

Technical Report AIP - 26

Mark Detweiler & Walter Schneider

Learning Research and Development Center
Univerity of Pittsburgh
Pittsburgh, PA 15260

29 September 1987

This research was supported by the Computer Sciences Division, Office of Naval Research and DARPA under Contract Numbers N00014-86-K-0678 and by ONR, Personnel and Training Research Programs, Psychological Sciences Division Contract Number N0014-86-K-0107; and by the Army Research Institute, under Contract No. MDA903-86-C-0149. Reproduction in whole or in part is permitted for purposes of the United States Government. Approved for public release; distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; Distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AIP 26	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Carnegie-Mellon University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Computer Sciences Division Office of Naval Research (Code 1133)
6c. ADDRESS (City, State, and ZIP Code) Department of Psychology Pittsburgh, Pennsylvania 15213	7b. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, Virginia 22217-5000
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Same as Monitoring Organization 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-K-0678
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS p400005ub2017-4-86 PROGRAM ELEMENT NO: N/A; PROJECT NO: N/A; TASK NO: N/A; WORK UNIT ACCESSION NO: N/A

11. TITLE (Include Security Classification)
A Connectionist/Control Architecture for Working Memory and Workload: Why Working Memory is Not 7+ /-2

12. PERSONAL AUTHOR(S)
M. Detweiler & W. Schneider

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM 86Sept15 TO 91Sept14	87 September 29	5

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Artificial Intelligence, connectionism, cognitive pschyology, working memory, long-term memory, short-term memory

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A runnable simulation architecture for working memory is described that provides an alternative to existing models of working memory, e.g., of Atkinson and Shiffrin (1968) and Baddeley (1986). It is used to interpret a variety of phenomena, including multiple resources, workload, chunking, sequential output, skilled and episodic memories, and stages of skill acquisition. The architecture is based on a set of modules organized into regions which communicate with each other on an innerloop of processing. A new feature of this architecture is a proposed context-storage module that temporarily stores context information in fast changing connection weights. This enables the system to expand effective working memory beyond the traditional 7±2 items. The context storage system is able to reload modules after short-term information decays or is displaced; in addition, it provides a means of achieving stable, robust processing under conditions of high workload.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Alan L. Meyrowitz	22b. TELEPHONE (Include Area Code) (202) 696-4302
	22c. OFFICE SYMBOL N00014

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

Proceedings of the Human Factors Society, 1987

A Connectionist/Control Architecture For Working Memory and Workload: Why Working Memory Is Not 7 +/- 2

Mark Detweiler & Walter Schneider

Learning Research and Development Center
University of Pittsburgh
3939 O'Hara St.
Pittsburgh, PA 15260

Abstract

A runnable simulation architecture for working memory is described that provides an alternative to existing models of working memory, e.g., of Atkinson & Shiffrin (1968) and Baddeley (1986). It is used to interpret a variety of phenomena, including multiple resources, workload, chunking, sequential output, skilled and episodic memories, and stages of skill acquisition. The architecture is based on a set of modules organized into regions which communicate with each other on an innerloop of processing. A new feature of this architecture is a proposed context-storage module that temporarily stores context information in fast changing connection weights. This enables the system to expand effective working memory beyond the traditional 7 +/- 2 items. The context storage system is able to reload modules after short-term information decays or is displaced; in addition, it provides a means of achieving stable, robust processing under conditions of high workload.

More than 30 years ago Miller (1956) described his persecution by an integer and proposed that the span of immediate memory could be best understood with respect to the "magic number seven, plus or minus two". In the intervening years, however, both Miller's characterization, as well as the modal model of short-term memory (e.g., Waugh & Norman, 1965; & Atkinson & Shiffrin, 1968, 1971) have proven to be inadequate models (cf. Crowder, 1982). For example, short-term memory can be both much smaller and larger than 7 +/- 2, e.g., being as small as size 2 in digit cancelling tasks and as large as 80+ digits in a span task (Chase & Ericsson, 1982). Furthermore, Baddeley and Hitch (1974) provide convincing evidence that short-term memory is much more complex and differentiated than previously supposed. In contrast to the unitary- and multiple-system theories of the late 1960s and early 1970s, they proposed the idea of a working memory system comprising separable, yet interacting subsystems. Since the publication of this classic paper, the concept of a working memory system has gained even greater support (see Baddeley, 1986).

The kinds of techniques that Baddeley and Hitch adopted required Ss to retain one or more items in short-term store while concurrently performing reasoning, language comprehension and learning tasks. According to the modal model, the concurrent memory loads should have occupied some of the 7 +/- 2 slots, and thus have adversely affected Ss' performance. Although holding six items did result in some degree of decrement on reasoning, e.g., showing a 3% to 14% increase in errors and a 44% slowing of the response, Ss could still perform the task. In addition, they found no effect on performance when only one or three items were preloaded. Hatano and Osawa (1983) found a similar lack of interference in the area of abacus calculation. When abacus experts were required to briefly rehearse a list of fruit names prior to the onset of the digit stream to be calculated, they showed only a slight drop in their digit spans. These results provide strong support for the idea that multiple short-term memory stores exist and that when material is stored in non-competing buffer-like memories, working memory capacity can be much larger than 7 +/- 2.

Despite the upsurge in interest in describing short-term memory as a working memory system, little progress has been made at illustrating how such a system might function within the context of a working simulation model. In what follows we briefly describe one architecture from a class of architectures which moves beyond box model descriptions and embodies a working simulation model capable of accounting for a range of working memory phenomena. A more detailed description can be found in Schneider and Detweiler (1987).

A Connectionist/Control Architecture Overview

The proposed architecture offers a means for representing a class of models of working memory and workload. It includes several different kinds of processing elements which result in mechanisms capable of achieving stable information processing of real-world

task components. It also provides a five-phase account of skill acquisition that is both gradual and continuous (see Schneider & Detweiler, 1987). Before outlining the three primary levels of detail, we briefly mention five assumptions which serve to constrain the architecture. First, we assume that processing occurs in *networks of* neural-like units organized into modules, sets of which in turn are organized into levels and regions. The regions, e.g., visual, speech, motor, etc. communicate with one another on an innerloop of connections. Second, we assume these modules process a restricted class of inputs, exhibiting local specialization of function. Third, we assume that the connection weights among the neural-like units store knowledge. Changes in knowledge due to learning are represented by the strength of the connections among units, or by the magnitudes of their weights. Fourth, we assume that connection weights may operate using a variety of rate constants. These constants determine how quickly the connections change as a function of intervening learning and length of the retention iterval. Finally, we assume information transferred within and between processing regions is modulated by a control processing system which controls information maintained in and output from the modules. This system is a variation of the CAP1 system used to account for automatic and controlled processing (see Shiffrin & Schneider, 1977; & Schneider & Mumme, in preparation). We next describe the architecture from three levels of scale.

There are three levels of scale at which to conceptualize and simulate the model. The first scale level, the **micro-level structure**, represents a neural-like network of processing units capable of producing associative processing and attentional phenomena. *Information is transferred between modules* via a message vector connecting output from one module as input to another module. The output, i.e., information flow, from a module is controlled by an attenuation unit within the module (see Schneider & Mumme, 1987; & Schneider & Detweiler, 1987 for details). The second scale level, the **macro-level structure**, represents interactions among a set of modules. Modules are organized in terms of *levels* and *regions* of processing. Levels represent successive processing stages, e.g., the auditory module consists of, at minimum, stages corresponding to phonemes, syllables, and words. The third scale level, the **system-level structure**, represents interactions among regions of modules (see Figure 1). Each region involves a series of levels of modules and corresponding control structures. At the system level there are two types of processing analogous to those at the macro level. The first is a *central control structure* which receives activity reports from all of the regions and is responsible for modulating message transmissions on what we refer to as the innerloop of processing. The second, the *innerloop*, represents the communication links among modules from regions possessing connections to one another.

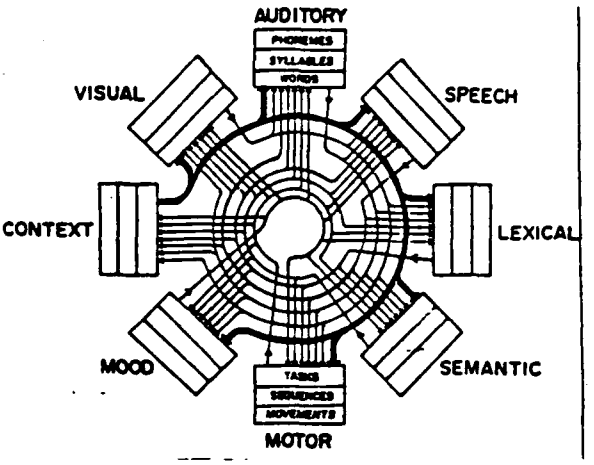

Figure 1. System-level model description. This diagram represents a top-down view of the processing regions found within the proposed architecture. Each region is divided into a series of bars to reflect a series of processing levels. A region's first or last level (last level of input regions and first level for output regions) is assumed to input to the innerloop of connections among regions. Modules on the innerloop have separate message vectors (pathways) to each of the other modules they connect to. All of the lines in Figure 1 represent message vectors. The context module transmits a message vector to all of the other modules on the innerloop. The context module's output is highlighted by a thick line to illustrate this connection pattern.

The central control structure receives control signals from each region and routes the signals among the various regions. For example, if the motor region requests an input, the central control structure may route the request to the speech transmitting module.

Here it is important to recognize that messages must *not* all pass through a central executive; instead, regions can communicate with other regions directly (see Figure 1). To minimize such interference from concurrent messages, a central control structure must sequence the region-based transmissions. This type of seriality can result in messages being delayed or omitted entirely, thus mimicing many of the item and order dropout phenomena seen in short-term memory literature and under conditions of high workload.

To model the robustness of human processing and match some of the learning data, the connectionist/control architecture embodies the notion of **context storage**. In a system in which considerable information must be maintained in working memory, context serves a special function. In this architecture a context-storage mechanism is proposed as a means of associating the

contents of messages on the innerloop to the temporal context in which they occur. In short, the context storage system is able to reload modules after short-term information has decayed or been displaced. It is also the mechanism responsible for maintaining stable, robust processing under conditions of high workload. To illustrate how context storage might work we have labeled one of the regions on the innerloop as context (see Figure 1). Context provides an association between the current time and what messages were transmitted by other modules of the innerloop. If a buffer is flushed by processing some interruption, the subject can activate the previous time context and revoke the previous contents of the buffer. If multiple traces are associated to the same context, interference occurs. Note, since context has connections to many modules (see Figure 1), it can provide a very large working memory.

Interpretations of Buffering

The connectionist/control architecture can be used to account for a number of phenomena found in the working memory and workload literatures. The first set of phenomena concern buffering. Three major experimental effects are often cited as supporting buffer models, namely, recency, decay, and span effects. The recency effect involves better recall of the last few items from a list (see Postman & Phillips, 1965). The decay effect is the exponential dropoff in probability of recall (Brown, 1958; Peterson & Peterson, 1959). The third phenomenon is a span effect found in digit and word span. Buffers are reflected in the connectionist/control architecture in the regional controllers and provide for the possibility of having them in multiple levels within the regional controllers as well. In order to sequentially input and output information in a buffering scheme requires techniques for clearing a module, maintaining a code within a module, and outputting a code from a module.

Working Memory Capacity

One of the misconceptions promoted by Miller's (1956) magic number seven and the modal models of short-term memory, e.g., Waugh & Norman (1965) and Atkinson & Shiffrin (1968, 1971) is that short-term memory capacity is limited to just a few items. Recently Card, Moran, and Newell (1983) argued that system/interface designers need to be able to analyze prospective tasks and to predict human performance on those tasks. They proposed a simplified human information processing model consisting of a set of goals, a set of operators, a set of methods for achieving goals, and a set of selection rules for choosing among competing methods and goals—commonly referred to as the GOMS model. Although the GOMS model is more sophisticated than variations of the modal model, e.g., replacing separate memory slots with slots that are subslots of each other, it fails to take into account the differential impact of different memory stores, rehearsing items early, and using context. In effect, the GOMS model offers a conservative view of working memory not well suited for making realistic predictions about a user's short-term memory capabilities, particularly when multiple memory stores are used concurrently.

To make informed predictions about working memory limitations it is important to begin to consider what often occurs in real-world tasks. For example, in electronic troubleshooting tasks a technician may be temporarily interrupted while tracing a fault, yet within a few seconds of resuming the task continue as if it had never occurred. An air-traffic controller many also be temporarily interrrupted yet gracefully resume the task. In general, as an operator is taxed with an ever greater mental load, his/her performance tends to become slower and somewhat more susceptible to errors, but there is no "red line" beyond which catastropic failure is inevitable.

Several results suggest that Miller's magic number should be called into question. Klapp, Marshburn, and Lester (1983) have clearly demonstrated that when subjects are allowed to briefly rehearse letters in a span task the resultant buffering does not interfere with their performance at judging the correctness of greater-than/less-than propositions (experiment #6), nor does it greatly interfere with performance on a variation of a Sternberg scanning task (experiment #7). Klapp and Philipoff (1983) also demonstated that it is possible to store and retain letters and concurrently process digits in a missing-digits task. Both of these sets of results run counter to the traditional view of short-term memory. Other experiments offer additional support for the idea of multiple buffers that can be strategically utilized to increase storage capacity without serious performance degradation in motor memory (Reisberg, Rappaport & O'Shaughnesy, 1984) and auditory visual digit span (Frick, 1984).

The results above are expected by, and can be accommodated within, the present architecture. In this model we assume many different buffers exist and suggest that the actual number available at a given level of processing will depend on the nature of the modules themselves. For example, auditory modules many be organized sequentially and visual modules spatially. We assume these buffers are limited in size and that the local control region can manage only a very small number of modules. The key feature of having may different buffers segmented into different processing classes is that by using several different classes simultaneously it is possible to enlarge the effective workspace of working memory well beyond 7 +/- 2. Moreover, the proposed context storage system provides an additional mechanism that can greatly increase the size of working memory storage. To appreciate the

magnitude of this increase, imagine one has R regions and N modules connecting to the innerloop from each region. The context vector could then be associated to (R-1)N modules. If there are 20 regions in the innerloop and 30 modules within each region that connect directly to the innerloop, then a single context vector and its association matrices could store 570 codes, i.e., one per module. Such a system offers short-term storage capacity far exceeding 7 +/- 2 items.

Mental Workload

The architecture presented here is well suited for investigating a number of important phenomena associated with the study of mental workload. It is able to simulate the performance of multiple concurrent tasks explicitly. First, as illustrated in Figure 1, the system encompasses a variety of resources which can be variously allocated to meet demands imposed by numerous task combinations (cf. Wickens, 1980). We assume these resources are limited such that when task demands exceed resource capacity, processing will be either delayed or errors will occur. Given this view of multiple resources, the connectionist/control architecture is consistent with views such as Wicken's (1980) that resources are differentiated by modality. These views divide, however, with respect to the nature of resource limitations. In addition to the resource-specific competition of multiple resource theory, this architecture accentuates the fundamental role played by competition for the central control structure.

There are at least five different strategies that can be invoked to perform concurrent tasks (see Schneider & Detweiler (1987) for details). *Strategy #1:* Buffer and delay messages for one of the tasks until the other task has been completed. *Strategy #2:* Redistribute task components into low-use buffers. *Strategy #3:* Use context storage to temporarily associate information to the present context and utilize it to load modules. Attending to context allows for the simultaneous loading of modules in numerous regions and initiating numerous concurrent processes. *Strategy #4:* Build automatic processes to reduce the processing load on the central and regional controllers. *Strategy #5:* Reduce message interference of concurrently transmitted messages. Interference can be reduced by decreasing the amount of time messages spend on the innerloop and by strengthening region-to-region connections on the innerloop.

Conclusion

Human working memory is both very limited (e.g., two digits in digit cancelling) and large (80 digits skilled memory), and this range is not interpretable within the modal buffer memory models (e.g., Atkinson & Shiffrin, 1968). The new connectionist/control architecture predicts the very limited working memory effects are the result of control processing limitations of the management of message transmission on an innerloop of processing or operating in a single region. In contrast, context storage provides a very large working memory as long as there are few associations of a given region to a given context.

Acknowledgement

This research was sponsored by the Army Research Institute, under Contract No. MDA903-86-C-0149 and Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research under Contract Nos. N-0014-86-K-0107 and N-00014-86-K-0678.

References

Atkinson, R.C., & Shiffrin, R.M. (1968). Human memory: A proposed system and its control process. In K.W. Spence & J.T. Spence (Eds.), The psychology of learning and motivation, Vol. 2. New York: Academic Press.

Atkinson, R.C., & Shiffrin, R.M. (1971). The control of short-term memory. Scientific American, 224, 82-90.

Baddeley, A.D. (1986). Working memory. Oxford: Clarendon Press.

Baddeley, A.D., & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), The psychology of learning and motivation, Vol. 8. New York: Academic Press.

Brown, J. (1958). Some tests of the decay theory of immediate memory. The Quarterly Journal of Experimental Psychology, 10, 12-21.

Card, S.K., Moran, T.P., & Newell, A. (1983). The psychology of human-computer interaction. Hillsdale, NJ: Erlbaum.

Chase, W.G., & Ericsson, K.A. (1982). Skill and working memory. In G.H. Bower (Ed.) The psychology of learning and motivation. Vol. 16. New York: Academic Press.

Crowder, R.G. (1982). The demise of short-term memory. Acta Psychologica, 50, 291-293.

Frick, R.W. (1984). Using both an auditory and a visual short-term store to increase digit span. Memory & Cognition, 12, 507-514.

Hatano, G., & Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. Cognition, 15, 95-110.

Klapp, S.T., Marshburn, E.A., & P.T. Lester (1983). Short-term memory does not involve the "working memory" of information processing: The demise of a common assumption. Journal of Experimental Psychology: General, 112, 240-264.

Klapp, S.T., & Philipoff, A. (1983). Short-term memory limits in performance. In A.T. Pope & L.D. Haugh (Eds.) Proceedings of the human factors society 27th annual meeting. Santa Monica, CA: Human Factors Society.

Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Peterson, L.R., & Peterson, J.J. (1959). Short-term retention of individual verbal items. Journal of Experimental Psychology, 58, 193-198.

Postman, L., & Phillips, L.W. (1965). Short-term retention of individual verbal items. Journal of Experimental Psychology, 70,, 473-478.

Reisberg, D., Rappaport, I., & O'Shaughnessy, M. (1984). Limits of working memory: The digit digit-span. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 203-221.

Schneider, W., & Detweiler, M. (1987). A connectionist/control architecture for working memory. In G.H. Bower (Ed.) The psychology of learning and motivation, Vol. 21. New York: Academic Press.

Schneider W., & Mumme, D. (1987). Attention, automaticity and the capturing of knowledge: A two-level cognitive architecture. Manuscript for submission.

Shiffrin, R.M., & Schneider, W. (1977). Controlled and automatic human information processing. II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84, 127-190.

Waugh, N.C., & Norman, D.A. (1965). Primary memory. Psychological Review, 72, 89-104.

Wickens, C.D. (1980). The structure of attentional resources. In R. Nickerson (Ed.). Attention and performance, VIII. Hillsdale, NJ: Erlbaum.